1 The Cantor Set

It is easy to see that \( C_n \) consists of \( 2^n \) disjoint closed intervals of length \( 3^{-n} \).

a) Prove that \( \mathcal{C} \) is compact and non-empty.

Since all the \( C_n \) are closed sets, and since an arbitrary intersection of closed sets is also closed, the set \( \mathcal{C} \) is closed. As it is a subset of \([0, 1]\) it is also bounded, hence compact. To see non-emptyness, note that clearly \( 0 \in \mathcal{C} \). In fact, it is easy to see that the endpoints of every closed interval \( C_{n,k} \) in the disjoint union \( C_n = C_{n,1} \cup C_{n,2} \cup \ldots \cup C_{n,2^n} \) belong to \( \mathcal{C} \).

b) Prove that \( \mathcal{C} \) is totally disconnected, i.e. given \( x \) and \( y \) in \( \mathcal{C} \) with \( x \neq y \) there is a \( x < z < y \) with \( z \notin \mathcal{C} \).

Given \( x < y \) we have \(|x - y| = \delta > 0 \). Choose \( n \) such that \( 3^{-n} < \delta \). Clearly \( x \) and \( y \) both have to be in \( C_n \). The length of each of the disjoint closed intervals in \( C_n \) is \( 3^{-n} \), so \( x \) and \( y \) have to lie in different connected components of \( C_n = C_{n,1} \cup C_{n,2} \cup \ldots \cup C_{n,2^n} \) (from left to right). If \( x \) lies in \( C_{n,k} \) then necessarily \( k < 2^n \) and the open interval between \( C_{n,k} \) and \( C_{n,k+1} \) contains only points from \( \mathcal{C} \).

c) Prove that \( \mathcal{C} \) does not have isolated points.

We need to show that any \( \delta \)-neighbourhood of an arbitrary \( x \in \mathcal{C} \) contains a point from \( \mathcal{C} \). Let hence \( x \in \mathcal{C} \) be given and \( \delta > 0 \) prescribed. Let \( n \) be such that \( 3^{-n} < \delta \). Clearly \( x \in C_n \), so \( x \) sits in one of the connected components of \( C_n = C_{n,1} \cup C_{n,2} \cup \ldots \cup C_{n,2^n} \), say \( C_{n,k} \). But since the left and right endpoints of the \( C_{n,k} \) are in \( \mathcal{C} \) and since \( C_{n,k} \) has length \( 3^{-n} \) we can find at least one point \( y \in \mathcal{C} \cap C_{n,k} \) such that \(|x - y| < 3^{-n} < \delta \).

d) Prove that \( m_*(\mathcal{C}) = 0 \).

Note that \( \mathcal{C} \subset C_n \) for any \( n \). By the monotonicity property of the exterior measure, we have \( m_*(\mathcal{C}) \leq m_*(C_n) \). Now since \( C_n \) is a disjoint union of \( 2^n \) compact intervals of length \( 3^{-n} \), we have \( m_*(C_n) = (2/3)^n \). We conclude that \( m_*(\mathcal{C}) < \epsilon \) for any \( \epsilon > 0 \) and hence \( m_*(\mathcal{C}) = 0 \).

e) Show that we can write \( C_n \) as

\[
C_n = \bigcup_{a_1, \ldots, a_n \in \{0, 2\}} \left[ \sum_{k=1}^{n} a_k 3^{-k}, \sum_{k=1}^{n} a_k 3^{-k} + \frac{1}{3^n} \right].
\]

We mentioned several times that \( C_n \) contains \( 2^n \) intervals of length \( 3^{-n} \). To show the claim of the hint, we first note that the left endpoint of \( C_n \) is 0 consistent with the formula. Assume now that for \( n = N \) the \( 2^N \) left endpoints of the closed disjoint intervals in \( C_N \) are indeed given by the \( 2^N \) numbers arising as \( \sum_{k=1}^{N} a_k 3^{-k} \) by different choices of \( a_k \in \{0, 2\} \). We now look at \( C_{N+1} \), which arises from \( C_N \) by deleting the mid-third intervals. Therefore, any of the \( 2^{2N} \) connected component of \( C_N \) (which each had 1 left
endpoints, say $p_i$) is decomposed into two different disjoint intervals with endpoints $p_i$ and $p_i + 23^{-N-1}$. Therefore, the left endpoints of $C_{N+1}$ are

$$\sum_{k=1}^{N} a_k 3^{-k} + 0 \quad \text{with} \quad a_k \in \{0, 2\} \quad \text{and} \quad \sum_{k=1}^{N} a_k 3^{-k} + 2 \cdot 3^{-N-1} \text{with} \quad a_k \in \{0, 2\}$$

But the collection of these points can be written as $\sum_{k=1}^{N+1} a_k 3^{-k}$ with $a_k \in \{0, 2\}$ and the proof by induction is completed. The overall claim then follows by noting that the intervals in $C_n$ have length $3^{-n}$.

f) Show that $x = \sum_{k=1}^{\infty} a_k 3^{-k}$ with $a_k \in \{0, 2\} \Leftrightarrow x \in \mathcal{C}$.

Let $x = \sum_{k=1}^{\infty} a_k 3^{-k}$ for some fixed collection of $(a_k)$. We need to show that $x \in C_n$ for any $n$. But for any $n$ we have

$$\sum_{k=1}^{n} a_k 3^{-k} \leq x \leq \sum_{k=1}^{n} a_k 3^{-k} + \sum_{k=1}^{\infty} 2 \cdot 3^{-k} = \sum_{k=1}^{n} a_k 3^{-k} + \frac{1}{3^n}$$

which is precisely the statement that $x$ is contained in one of the intervals of $C_n$.

Conversely, let $x \in \mathcal{C}$, so $x$ is in any $C_n$. We construct a (unique) sequence $(a_k)$ (consisting of 0’s and 2) such that we have

$$\sum_{k=1}^{n} a_k 3^{-k} \leq x \leq \sum_{k=1}^{n} a_k 3^{-k} + \frac{1}{3^n}. \quad (1)$$

for any $n$ as follows. Clearly for any fixed $N$ there are unique $a_k$ with $k = 1, \ldots, N$ satisfying (1) with $n = N$ as $x$ lies in exactly one of the disjoint intervals of $C_N$. Similarly, since $x \in C_{N+1}$, there are unique $\tilde{a}_k$ with $k = 1, \ldots, N + 1$ satisfying (1) with $n = N + 1$. We now show that $a_k = \tilde{a}_k$ for $k = 1, \ldots, n$, which (by induction) shows that the the coefficients $a_k$ in (1) do not depend on $n$. To see this, note that if $x \in C_n$, then the interval of $C_{n+1}$ of which $x$ is an element can only be the left third or the right third of the interval of $C_n$ of which $x$ is an element of. So either

$$\sum_{k=1}^{N} a_k 3^{-k} + \frac{0}{3^{N+1}} \leq x \leq \sum_{k=1}^{N} a_k 3^{-k} + \frac{1}{3^{N+1}} \quad \text{or} \quad \sum_{k=1}^{N} a_k 3^{-k} + \frac{2}{3^{N+1}} \leq x \leq \sum_{k=1}^{N} a_k 3^{-k} + \frac{3}{3^{N+1}}.$$

In the first case,

$$\sum_{k=1}^{N+1} a_k 3^{-k} \leq x \leq \sum_{k=1}^{N+1} a_k 3^{-k} + \frac{1}{3^{N+1}}$$

holds with $a_{k+1} = 0$ while in the second the above holds with $a_{k+1} = 2$. This shows in particular, $a_k = \tilde{a}_k$ for $k = 1, \ldots, n$ as desired.

With (1) established for all $n$ we observe that the sum on the left converges and that in fact the left hand side and the right hand side converge to the same value, giving $x = \sum_{k=1}^{\infty} a_k 3^{-k}$ as claimed.

g) Define the Cantor-Lebesgue function $F : \mathcal{C} \to [0, 1]$ as

$$F(x) = \sum_{k=1}^{\infty} \frac{b_k}{2^k} \quad \text{for} \quad x = \sum_{k=1}^{\infty} a_k 3^{-k}, \text{where} \quad b_k = \frac{a_k}{2}.$$ 

Show that $F$ is well-defined and in fact continuous on $\mathcal{C}$. Show also that $F$ is surjective. Conclude that $\mathcal{C}$ is uncountable.

To show that $F$ is well defined we need to show that the $(a_k)$ in the expansion $x = \sum_{k=1}^{\infty} a_k 3^{-k}$ are unique. This is easy to see because suppose there were two sequences $(a_k)$ and $(\tilde{a}_k)$ giving rise to the same $x$ and differing at position $N$ where $a_N = 0$ and $\tilde{a}_N = 2$ (or the other way round). Then we must have

$$0 + \sum_{N+1}^{\infty} a_k 3^{-k} = \frac{2}{3^N} + \sum_{N+1}^{\infty} \tilde{a}_k 3^{-k}$$
and in particular
\[ 0 + \sum_{N+1}^{\infty} 2 \cdot 3^{-k} \geq \frac{2}{3^N} \] which implies \( \frac{1}{3^N} \geq \frac{2}{3^N} \), a contradiction.

Surjectivity follows from the fact that every real number has a (non-unique) dyadic expansion. Surjectivity implies that the cardinality of \( C \) is at least that of \([0,1]\) but since the latter is a subset of the former, they must have the same cardinality.

To show continuity at \( x \in C \), we let \((x_n)\) be an arbitrary sequence in \( C \) with \( x_n \to x \). We need to show that given \( \epsilon > 0 \) we can find \( N \) such that for all \( n \geq N \) we have \( |F(x) - F(x_n)| < \epsilon \).

We fix a \( K \) large such that \( \sum_{k=K+1}^{\infty} \frac{1}{2^k} < \epsilon \). We let \( x_n = \sum_{k=1}^{\infty} (a_n)_k 3^{-k} \) and \( x = \sum_{k=1}^{\infty} a_k 3^{-k} \). (We already know that the \((a_n)_k \) and \( a_k \) are unique.) Since we have \( x_n \to x \) we can choose an \( N \) such that we have \( |x - x_n| < \frac{1}{2^{K+1}} \) for all \( n \geq N \). It is easy to see that this implies \((a_n)_k = a_k \) for all \( k = 1, \ldots, K \) and \( n \geq N \) (get a contradiction using geometric series). With these choices, we have for all \( n \geq N \)
\[ |F(x_n) - F(x)| = \left| \sum_{k=1}^{\infty} \frac{(a_n)_k - a_k}{2} \right| = \left| \sum_{K+1}^{\infty} \frac{(a_n)_k - a_k}{2} \right| \leq \sum_{K+1}^{\infty} \frac{1}{2^k} < \epsilon \]
as desired. Remark: \( F \) can actually be extended to a continuous function on all of \([0,1]\).

2 Fat Cantor Sets

a) Show that \( m_*(\hat{C}) = \frac{1}{2} \) and conclude that \( \hat{C} \) is uncountable.

Note that \( \hat{C} \) is measurable and \( m(\hat{C}) = m_*(\hat{C}) \). It is also clear that \( \hat{C}_k \) is measurable and we compute the measure of its complement in \([0,1]\) (the intervals we remove) for \( k \geq 1 \) to be
\[ \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots + \frac{1}{2^{2j}} = \frac{1}{2} \sum_{j=1}^{k} \frac{1}{2^j} = \frac{1}{2} - \frac{1}{2^{k+1}} \cdot \]

Hence \( m(\hat{C}_k) = \frac{1}{2} + \frac{1}{2^{k+1}} \). We have \( \hat{C}_k \supseteq \hat{C}_{k+1} \) and \( m(\hat{C}_1) = 1 \) and hence by a result from lectures \( m(\bigcap_{k=1}^{\infty} \hat{C}_k) = \lim_{k \to \infty} m(\hat{C}_k) = \frac{1}{2} \). If \( \hat{C} \) was countable we would necessarily have \( m_*(C) = 0 \).

b) Show that \( \hat{C} \) is again compact, totally disconnected and has no isolated points.

The proof is very similar to the first question and will not be repeated.

3 *Characterisation of Riemann integrable functions

Prove that a bounded function on an interval \([a,b]\) is Riemann integrable if and only if its set of discontinuities has measure zero.

This is a hard problem. You should look up the outline of the proof in Stein-Shakarchi (p.47) or elsewhere in the literature. You can use the result at the end of Exercise 4.

4 Limits of continuous functions \( f : [0,1] \to \mathbb{R} \) (Exercise 10 of \([SS]\))

In this exercise we construct the sequence of continuous functions promised in the first lecture.

Let \( \hat{C} \) denote the Fat Cantor Set constructed in Exercise 2. We define the function \( F_1 \) to be a piecewise linear continuous function which is equal to 1 on the complement of the first interval \( I_1 = [3/8, 5/8] \) removed
from \([0,1]\) in the construction of \(\hat{C}\) and zero at the point at the centre of \(I_1\). Similarly we construct \(F_2\) to be a piecewise linear continuous function which is equal to 1 on the complement of the open intervals removed at the second stage and zero at the centre of these intervals. Continuing in this way we can define \(F_n\) for all \(n\) and finally

\[ f_n = F_1 \cdot F_2 \cdots F_n \]

a) Show that for all \(n \geq 1\) and all \(x \in [0,1]\) one has \(0 \leq f_n(x) \leq 1\) and \(f_n(x) \geq f_{n+1}(x)\). We conclude that \(f_n(x)\) converges to a limit which we denote by \(f(x)\).

All easy to see.

b) Show that the function \(f\) is discontinuous at every point of \(\hat{C}\). Conclude that \(f\) is not Riemann integrable (despite the sequence \(s_n = \int f_n\) converging).

Hint: Note that \(f(x) = 1\) if \(x \in \hat{C}\) and find a sequence of points \(\{x_n\}\) so that \(x_n \to x\) and \(f(x_n) = 0\).

The last conclusion follows from Exercise 3 so we will focus on establishing discontinuity at the points \(x \in \hat{C}\). We first define a sequence \(x_n\) with \(|x_n - x| < 2^{-n}\) as follows. We know that each \(\hat{C}_k\) consists of \(2^k\) disjoint intervals of length \(L_k\) necessarily smaller than \(2^{-k}\) (the actual length can be computed to be \(L_k = 2^k + 1\)). Therefore, given \(n\) we know that \(x\) is contained in one of the intervals of \(\hat{C}_n\) of length \(L_n\). At the next stage, \(L_n\) is tri-sected and an interval of length \(4^n + 1\) is removed from the middle. We choose \(x_n\) to be the point in the middle of that interval. Clearly \(|x - x_n| < L_n < 2^{-n}\) by construction. Moreover, since \(x_n\) sits at the middle point of an interval that gets removed, we have \(F_{n+1}(x_n) = 0\) for all \(n\) and hence \(\lim_{k \to \infty} f_k(x_n) = f(x_n) = 0\) for any \(n\). This is the desired sequence which clearly proves that \(f\) is not continuous at \(x\).

5 The outer Jordan content of a set (Exercise 14 of [SS])

The outer Jordan content \(J_\star(E)\) of a bounded set \(E\) in \(\mathbb{R}\) is defined by

\[ J_\star(E) = \inf \sum_{j=1}^{N} |I_j| \]

where the infimum is taken over all finite coverings \(E \subset \bigcup_{j=1}^{N} I_j\) by intervals \(I_j\).

a) Prove that \(J_\star(E) = J_\star(\overline{E})\) for every set \(E\). Here \(\overline{E}\) denotes the closure of \(E\).

Wlog we can assume that the intervals \(I_j\) of the finite covering are closed (why?).

But then, if \(E\) is covered by finitely many closed intervals \(I_i\) its closure \(\overline{E}\) is also covered by those intervals because the closure is the smallest closed set containg \(E\) and the union of the finitely many closed intervals is a closed set containing \(E\). It follows that \(J_\star(E) \geq J_\star(\overline{E})\). Since the other direction is clear by monotonicity, the claim follows.

b) Exhibit a countable subset \(E \subset [0,1]\) such that \(J_\star(E) = 1\) while \(m_\star(E) = 0\).

Take \(E = \mathbb{Q} \cap [0,1] \subset [0,1]\). Since the rationals are dense in \([0,1]\), their closure is all of \([0,1]\) which has outer Jordan content 1. But \(E\) is countable, so \(m_\star(E) = 0\).